

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **06236197 A**

(43) Date of publication of application: **23.08.94**

(51) Int. Cl.

G10L 3/02
G10L 3/00

(21) Application number: **05180056**

(22) Date of filing: **24.06.93**

(30) Priority: **30.07.92 JP 04223474**
16.12.92 JP 04354513

(71) Applicant: **RICOH CO LTD**

(72) Inventor: **SAKAYORI TETSUYA**

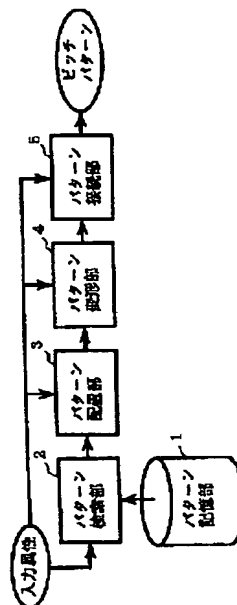
(54) PITCH PATTERN GENERATION DEVICE

(57) Abstract:

PURPOSE: To attain generating a pitch pattern applicable to a device synthesizing an optional voice like regular synthesis and being natural more than usual.

CONSTITUTION: This device is constituted so that a unit pattern cut out from the pitch pattern of human vocalization with a proper unit is stored in a pattern storage part 1, and the unit pattern suited to an input attribute is retrieved from the pattern storage part 1 based on the prescribed input attribute in a pattern retrieval part 2, and the unit pattern obtained by the retrieval is arranged on a time base in a pattern arrangement part 3, and the unit pattern arranged on the time base is deformed and the pitch pattern is generated in a pattern deformation part 4, and since the pitch pattern vocalized by a human is used as it is, the pitch pattern capable of expressing natural change and fluctuation, nuance, etc., being hardly generated with a rule is generated.

COPYRIGHT: (C)1994,JPO&Japio



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平6-236197

(43) 公開日 平成6年(1994)8月23日

(51) Int. Cl. ⁵

G10L 3/02

3/00

識別記号

庁内整理番号

F I

技術表示箇所

D 8946-5H

H 8946-5H

審査請求 未請求 請求項の数26 F D (全20頁)

(21) 出願番号 特願平5-180056

(22) 出願日 平成5年(1993)6月24日

(31) 優先権主張番号 特願平4-223474

(32) 優先日 平4(1992)7月30日

(33) 優先権主張国 日本 (J P)

(31) 優先権主張番号 特願平4-354513

(32) 優先日 平4(1992)12月16日

(33) 優先権主張国 日本 (J P)

(71) 出願人 000006747

株式会社リコー

東京都大田区中馬込1丁目3番6号

(72) 発明者 酒寄 哲也

東京都大田区中馬込1丁目3番6号 株式

会社リコー内

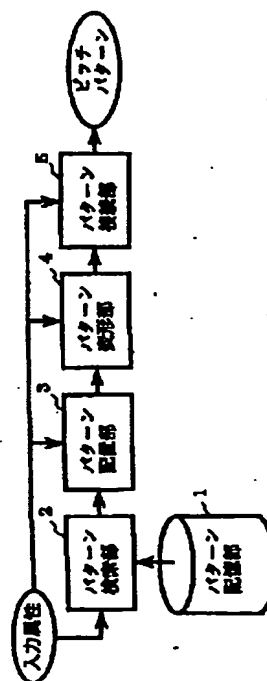
(74) 代理人 弁理士 植本 雅治

(54) 【発明の名称】 ピッチパターン生成装置

(57) 【要約】

【目的】 規則合成のような任意の音声を合成する装置への適用に適し、従来に比べてより自然なピッチパターンを生成可能である。

【構成】 人間の発声のピッチパターンから適当な単位で切り出したユニットパターンをパターン記憶部1に記憶し、パターン検索部2では、所定の入力属性に基づき、該入力属性に適合したユニットパターンをパターン記憶部1から検索し、パターン配置部3では、検索により得られたユニットパターンを時間軸上に配置し、パターン変形部4では、時間軸上に配置されたユニットパターンを変形して、ピッチパターンを生成するようにしており、人間の発声したピッチパターンをそのまま使用しているので、規則による生成では困難な自然な変化やゆらぎ、ニュアンスなどを表現可能なピッチパターンを生成できる。



【特許請求の範囲】

【請求項1】 人間の発声のピッチパターンから適当な単位で切り出したユニットパターンを記憶するパターン記憶手段と、所定の入力属性または入力言語情報に基づき、該入力属性または入力言語情報に適合したユニットパターンを検索するパターン検索手段と、検索により得られたユニットパターンを時間軸上に配置するパターン配置手段と、時間軸上に配置されたユニットパターンを変形するパターン変形手段とを有していることを特徴とするピッチパターン生成装置。

【請求項2】 請求項1記載のピッチパターン生成装置において、前記ユニットパターンとしては、アクセント句の上昇部と下降部とをそれぞれ1つ含むものを1つの単位として切り出したパターン、アクセント句の上昇部または下降部を1つ含むものを1つの単位として切り出したパターン、アクセント句の平坦部を含むピッチパターン区間を1つの単位として切り出したパターン、フレーズ末の特徴的ピッチ変化を示す部分を1つの単位として切り出したパターン、特定の範囲の変化率を示すピッチパターン区間を1つの単位として切り出したパターン、特定の範囲の曲率を示すピッチパターン区間を1つの単位として切り出したパターンのうちの、少なくとも1つが用いられることを特徴とするピッチパターン生成装置。

【請求項3】 請求項1記載のピッチパターン生成装置において、前記パターン記憶手段には、切り出されたユニットパターンが記憶されるか、あるいは切り出されたユニットパターンをベクトル量子化したものが記憶されるようになっていることを特徴とするピッチパターン生成装置。

【請求項4】 請求項1記載のピッチパターン生成装置において、前記パターン検索手段は、入力属性として、ピッチ変化部分近傍の複数モーラにわたる音韻環境情報、ユニットパターン内あるいはその近傍におけるプロミネンスに関する情報、ユニットパターンが含まれるフレーズのイントネーションスタイルの情報のうちの少なくとも1つの情報を使用するようになっていることを特徴とするピッチパターン生成装置。

【請求項5】 請求項1記載のピッチパターン生成装置において、前記パターン検索手段は、入力属性からニューラルネットワークによって最適なユニットパターンを決定するようになっていることを特徴とするピッチパターン生成装置。

【請求項6】 請求項1記載のピッチパターン生成装置において、前記パターン配置手段は、アクセント句境界の時間軸上の位置、アクセント核の時間軸上の位置、フレーズ末の音韻、音節、モーラなどの境界の時間軸上の位置、ピッチ変化部分近傍の音韻環境情報のうちの少なくとも1つを入力情報としてユニットパターンの時間軸上の位置を決定するようになっていることを特徴とする

ピッチパターン生成装置。

【請求項7】 請求項1記載のピッチパターン生成装置において、前記パターン変形手段は、ユニットパターンがアクセント句内の平坦部境界情報を持つ場合には、平坦部を時間軸方向に伸縮するような変形を施し、ユニットパターンがアクセント句内の上昇部と下降部の境界情報を持つ場合には、上昇部と下降部をピッチ軸方向に伸縮するような変形を施すようになっていることを特徴とするピッチパターン生成装置。

10 【請求項8】 請求項1記載のピッチパターン生成装置において、前記パターン変形手段によって変形されたユニットパターンを接続して一連のフレーズのパターンを生成するパターン接続手段がさらに設けられていることを特徴とするピッチパターン生成装置。

【請求項9】 請求項8記載のピッチパターン生成装置において、前記パターン接続手段は、ユニットパターン間を直線で補間して接続するか、あるいは、ユニットパターン間を3次曲線で補間して接続するようになっていることを特徴とするピッチパターン生成装置。

20 【請求項10】 請求項1記載のピッチパターン生成装置において、前記パターン記憶手段に記憶されるユニットパターンとしては、特定の言語状況において特徴的に現われるピッチパターンが用いられることを特徴とするピッチパターン生成装置。

【請求項11】 請求項10記載のピッチパターン生成装置において、前記ユニットパターンの1つとして、呼吸段落や文などの、一連のフレーズにおける特定の位置において特徴的に現われるピッチパターンが用いられることを特徴とするピッチパターン生成装置。

30 【請求項12】 請求項10記載のピッチパターン生成装置において、前記ユニットパターンの1つとして、頭高型、平板型などの特定のアクセント型において特徴的に現われるピッチパターンが用いられることを特徴とするピッチパターン生成装置。

【請求項13】 請求項10記載のピッチパターン生成装置において、前記ユニットパターンの1つとして、プロミネンスが付与されている部分、副次アクセント核が現われる部分などの特定のアクセントレベルにおいて特徴的に現われるピッチパターンが用いられることを特徴とするピッチパターン生成装置。

40 【請求項14】 請求項10記載のピッチパターン生成装置において、前記ユニットパターンの1つとして、間投詞や終助詞、擬声語、擬態語、掛け声などのその語に特有のピッチパターンで発声される語が用いられるか、あるいはフレーズのピッチパターンがそのまま用いられることを特徴とするピッチパターン生成装置。

【請求項15】 請求項10記載のピッチパターン生成装置において、前記ユニットパターンの1つとして、数
50 字列、韻文などに周期的に現われる特徴的ピッチパターンが用いられることを特徴とするピッチパターン生成装

置。

【請求項16】 請求項10記載のピッチパターン生成装置において、前記ユニットパターンの1つとして、語や音韻から独立して主に韻律によって特定の意図を表わす特徴的ピッチパターンが用いられることを特徴とするピッチパターン生成装置。

【請求項17】 請求項1記載のピッチパターン生成装置において、前記パターン記憶手段には、折れ線近似されたユニットパターンが記憶されることを特徴とするピッチパターン生成装置。

【請求項18】 請求項17記載のピッチパターン生成装置において、前記折れ線近似されたユニットパターンは、始点のピッチおよび各折れ線のそれぞれの区間と折れ線の傾きとの表現データによって表現されて前記パターン記憶手段に記憶されることを特徴とするピッチパターン生成装置。

【請求項19】 請求項18記載のピッチパターン生成装置において、各折れ線の傾きは、ピッチ周期が単位長だけ変化するのに要する時間としてのステップ値によって表現されており、この場合、前記パターン検索手段は、パターン記憶手段に記憶されているユニットパターンの表現データを検索し、該表現データに基づきユニットパターンを再生するときに、ステップ値が正の値の場合、ピッチ周期を単位長増加させ、ステップ値が負の値の場合、ピッチ周期を単位長減少させる処理を行なうようになっていることを特徴とするピッチパターン生成装置。

【請求項20】 請求項19記載のピッチパターン生成装置において、前記パターン変形手段は、各折れ線の傾きを表わすステップ値に一定のバイアスを加えることによって、前記ユニットパターンにピッチ方向の滑らかな変形を施すことを特徴とするピッチパターン生成装置。

【請求項21】 請求項20記載のピッチパターン生成装置において、前記パターン変形手段は、ピッチ更新周期を“2”の補数表現によって扱い、バイアスを加えた結果のステップ値の絶対値が最大値を越えた場合に、これを反対符号の最大絶対値とすることを特徴とするピッチパターン生成装置。

【請求項22】 請求項20または請求項21記載のピッチパターン生成装置において、ピッチ変化量とバイアスとの対応関係が予め求められている場合に、前記パターン変形手段は、該対応関係を用いてユニットパターンのピッチ方向の変形を行なうことを特徴とするピッチパターン生成装置。

【請求項23】 請求項1記載のピッチパターン生成装置において、さらに、前記パターン記憶手段に記憶されているユニットパターンと対応付けられて、条件を含むルールが記憶されるルール記憶手段が設けられており、前記パターン検索手段は、モーラなどの入力言語情報単

位ごとに前記ルール記憶手段に記憶されているルールの条件を調べることによって、ルールの条件を満たすユニットパターンを前記パターン記憶手段から検索するようになっていることを特徴とするピッチパターン生成装置。

【請求項24】 請求項23記載のピッチパターン生成装置において、前記ルールの条件には、言語情報単位の特徴を表わす複数の変数の値の範囲が含まれており、前記パターン検索手段は、ルールの条件に含まれる全ての変数について、入力言語情報単位の値がその範囲内にあるとき当該ルールを採用することによって、ユニットパターンを検索するようになっていることを特徴とするピッチパターン生成装置。

【請求項25】 請求項24記載のピッチパターン生成装置において、ルールの条件に含まれる変数として、当該言語情報単位の句頭あるいは句末からの位置を用いるか、または、当該言語情報単位を含む句の隣接句との境界種別を用いるか、または、当該言語情報単位を含む句のアクセント情報を用いるか、または、当該言語情報単位の音韻情報を用いることを特徴とするピッチパターン生成装置。

【請求項26】 請求項23記載のピッチパターン生成装置において、1つのルールに複数のユニットパターンが対応付けられており、ルールの条件を満たすユニットパターンが複数検索される場合に、ゆれを与えて1つのユニットパターンを選択することを特徴とするピッチパターン生成装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、規則音声合成装置などに用いられ、ピッチ周波数の時間パターン、すなわちピッチパターンを生成するピッチパターン生成装置に関する。

【0002】

【従来の技術】 従来、規則音声合成装置による合成音は、人間の発声とは明らかに異なる機械的で単調なものであり、自然さに欠け、聴く者にとって違和感があった。自然音声に近い合成音を得るため、近年、例えば特開平2-197897号（以下、従来例1と称す）、あるいは特開平3-139699号（以下、従来例2と称す）に示されているような装置が提案されている。

【0003】 すなわち、従来例1には、単語あるいは文節のアクセント型に応じてピッチ周波数が上昇するモーラにおける音素の種類ごとにピッチ周波数の上昇パターンを記憶し、また、単語あるいは文節のアクセント型に応じてピッチ周波数が下降するモーラにおける音素の種類ごとにピッチ周波数の下降パターンを記憶し、合成する単語あるいは文節のアクセント型に応じてピッチ周波数が上昇するモーラにおける音素の種類に応じて上記記憶されたピッチ周波数の上昇パターンを読み出し、合成

10

20

30

40

50

する単語あるいは文節のアクセント型に応じてピッチ周波数が下降するモーラにおける音素の種類に応じて上記記憶されたピッチ周波数の下降パターンを読み出し、読み出されたピッチ周波数の上昇パターンと読み出されたピッチ周波数の下降パターンとを接続して単語あるいは文節全体のピッチパターンを生成するピッチパターン生成装置が提案されており、この装置では、ピッチ周波数レベルが変化するモーラにおいて、自然音声のピッチパターンに近いピッチパターンを生成することを意図している。

【0004】また、従来例2には、韻律語を単位として音声を記憶し、かつピッチパターンが任意に制御できる音声編集合成器において、各韻律語の記憶時にピッチについてはアクセント成分のみを記憶しておき、文章合成時にその文章に対応したフレーズ成分を生成するとともに、該フレーズ成分にアクセント成分を重畳することによりピッチパターンを生成する音声編集合成器が提案されており、この音声編集合成器では、文章全体としての抑揚が自然となるようなピッチパターンを生成することを意図している。

【0005】

【発明が解決しようとする課題】このように、従来例1および従来例2に示されているいずれの装置も、自然音声に近い合成音を得ることを意図している。しかしながら、従来例1のピッチパターン生成装置では、使用するパターンはアクセント型に応じてピッチが上昇あるいは下降する部分だけに限られているので、そのほかの部分の不自然性は解消されない。また、このパターンはピッチ周波数が「高」から「低」へ、あるいはその逆のレベルへ変化する位置の音素の種類だけによって分類されているので、同じ音韻で変化する場合にはいつも同じパターンが使用され単調な音声となり、さらに、周囲の音韻環境やプロミネンスの有無などが考慮されていないため、不自然さが残るという問題があった。

【0006】また、従来例2の装置は、音声編集合成器に関するものであるので、使用する全ての韻律語に対応してアクセント成分を記憶している必要があるが、韻律語は一般に複数の単語や文節の組み合わせからなるため、全ての韻律語を網羅して記憶することは現実的に不可能である。また、文章内の位置、文脈、イントネーション、プロミネンス、発話速度などが変化した場合にはそのままでは使用できない。従って、規則合成のような任意の音声を合成する装置への適用は難しいという問題があった。

【0007】本発明は、このような問題を解決し、規則合成のような任意の音声を合成する装置への適用にも適しており、従来に比べてより自然なピッチパターンを生成するピッチパターン生成装置を提供することを目的としている。

【0008】

【課題を解決するための手段および作用】上記目的を達成するために、請求項1記載の発明は、人間の発声のピッチパターンから適当な単位で切り出したユニットパターンを記憶するパターン記憶手段と、所定の入力属性に基づき、該入力属性に適合したユニットパターンを検索するパターン検索手段と、検索により得られたユニットパターンを時間軸上に配置するパターン配置手段と、時間軸上に配置されたユニットパターンを変形するパターン変形手段とを有している。このように、人間の発声したピッチパターンをそのまま使用することによって、規則による生成では困難な自然な変化やゆらぎ、ニュアンスなどを表現可能なピッチパターンを生成することができる。

【0009】また、請求項2記載の発明において、アクセント句の上昇部と下降部とをそれぞれ1つ含むものを1つの単位として切り出したものをユニットパターンとして使用する場合には、ピッチパターンの重要な単位であるアクセント句のパターンをほぼ保存したピッチパターンを生成することができる。また、アクセント句の上昇部または下降部を1つ含むものをユニットパターンとして使用する場合には、簡単な接続処理によってピッチパターンを生成することができる。また、アクセント句の平坦部を含むユニットパターンを使用する場合には、定常部のゆらぎ等を保存でき、自然なピッチパターンを生成することができる。また、フレーズ末の特徴的ピッチ変化をユニットパターンとして使用する場合には、規則による生成では困難な多彩なイントネーションを表現することができる。また、特定の範囲のピッチ変化率を示すパターンをユニットパターンとして使用する場合には、変化部分の概形や定常部分のゆらぎ等を保存することができ、規則による生成では困難であった多彩で自然なピッチパターンを生成することができる。また、特定の範囲のピッチ曲率を示すパターンをユニットパターンとして使用する場合には、その他の部分を直線で補間できるため、処理量を低減することができる。

【0010】また、請求項3記載の発明において、ベクトル量子化されたユニットパターンを記憶し、これを使用するときには、パターン記憶手段の容量を大幅に削減することができ、装置の低コスト化、小型化、処理の高速化を図ることができる。

【0011】また、請求項4記載の発明において、ピッチ変化部分近傍の複数モーラにわたる音韻環境からユニットパターンを検索する場合には、音韻環境によるパターンの局所的变化の再現性が向上し、より緻密なピッチパターンを生成することができる。また、プロミネンスの情報からユニットパターンを検索する場合には、従来単調になりがちだった合成音声の抑揚にめりはりがつき、了解性、自然性を向上させることができる。また、イントネーションスタイルからユニットパターンを検索する場合には、規則による生成では表現が困難なイント

ネーションを正確に再現することができる。

【0012】また、請求項5記載の発明は、ニューラルネットワークによってユニットパターンを検索するようになっており、これにより、学習した人間の発声パターンに近いパターンを再現することができ、自然性を向上させることができる。

【0013】また、請求項6記載の発明において、パターン配置手段により、アクセント句境界の位置を用いてユニットパターンを配置する場合には、アクセント句の上昇開始点をほぼ再現でき、了解性を向上させることができる。また、アクセント核の位置を用いてユニットパターンを配置する場合には、アクセント句の下降開始点をほぼ再現でき、了解性を向上させることができる。また、フレーズ末の音韻、音節、モーラなどの境界位置を用いてユニットパターンを配置する場合には、イントネーションを効果的に表現でき、自然性を向上させることができる。また、ピッチ変化部分近傍音韻環境情報を用いてユニットパターンを配置することにより、ピッチパターンの時間構造をより緻密に再現でき、自然性を向上させることができる。

【0014】また、請求項7記載の発明では、パターン変形手段は、アクセント句の平坦部を時間軸方向に伸縮するようにしており、ピッチ変化部については時間軸方向に変形しないことにより、ピッチ変化部の時間構造を保存することができ、安定した自然なピッチパターンを生成することができる。また、アクセント句の上昇部と下降部をピッチ軸方向の変形をするようにしており、平坦部についてはピッチ軸方向に変形しないことにより、ピッチ定常部部の高さやゆらぎ等を保存することができ、プロミネンスなども含めたパターン表現が可能となり自然性を向上させることができる。

【0015】また、請求項8、9記載の発明では、パターン変形手段によって変形されたユニットパターンを接続して一連のフレーズのパターンを生成するパターン接続手段がさらに設けられており、パターン接続手段において、ユニットパターン間を直線で補間する場合には、処理量を低減できて、高速化や低コスト化を図ることができ、また、ユニットパターン間を3次曲線で補間する場合には、人間のピッチ変化に似た概形を表現でき、自然なピッチパターンを生成することができる。

【0016】また、請求項10乃至16記載の発明では、パターン記憶手段に記憶されるユニットパターンとして、特定の言語的状況において特徴的に現われるピッチパターンが用いられることを特徴としている。これにより、規則音声合成の入力情報である言語情報から、ヒューリスティックにパターンを検索することができ、効率的で効果的なピッチパターン生成を実現することができる。

【0017】また、請求項17、請求項18記載の発明では、パターン記憶手段には、折れ線近似されたユニッ

トパターンが記憶されるようになっており、この際、折れ線近似されたユニットパターンは、始点のピッチおよび各折れ線のそれぞれの区間と折れ線の傾きとの表現データによって表現されてパターン記憶手段に記憶されることにより、ユニットパターンのデータ量を減らし、メモリ量と処理量を軽減でき、低コスト化、高速化を図ることができる。さらに、ユニットパターンのピッチ方向のシフトは始点ピッチの増減のみで、時間方向の伸縮は区間長の増減によって、それぞれ簡単に実現可能で、ユニットパターンの変形が簡単な処理によって実現できるため、高速化、高機能化を図ることができる。

【0018】また、請求項19記載の発明では、上記各折れ線の傾きは、ピッチ周期が単位長だけ変化するのに要する時間としてのステップ値によって表現されており、この場合、パターン検索手段は、パターン記憶手段に記憶されているユニットパターンの表現データを検索し、該表現データに基づきユニットパターンを再生するときに、ステップ値が正の値の場合、ピッチ周期を単位長増加させ、ステップ値が負の値の場合、ピッチ周期を単位長減少させる処理を行なうようになっている。これにより、カウンタ、比較器などによる非常に簡単な構成で折れ線近似のユニットパターンを再生し、実際のピッチパターンを得ることができる。

【0019】また、請求項20記載の発明では、パターン変形手段は、各折れ線の傾きを表わすステップ値に一定のバイアスを加えることによって、ユニットパターンにピッチ方向の滑らかな変形を施す。これにより、ユニットパターン間の接続やプロミネンス付与などで必要となるパターン変形処理が加算器だけの非常に簡単な構成で実現できる。

【0020】また、請求項21記載の発明では、パターン変形手段は、ステップ値を“2”の補数表現によって扱い、バイアスを加えた結果のステップ値の絶対値が最大値を越えた場合に、これを反対符号の最大絶対値とする。これにより、右上がりのパターンから右下がりのパターンへ、またはその逆へ、連続的にかつ無理なく、パターン変形する処理を簡単に実現することができる。

【0021】また、請求項22記載の発明では、ピッチ変化量とバイアスとの対応関係が予め求められている場合に、パターン変形手段は、該対応関係を用いてユニットパターンのピッチ方向の変形を行なう。これにより、ユニットパターン間の接続やプロミネンス付与などで必要となるバイアスを、処理量の多い乗除算を必要とすることなく、極めて容易に求めることができ、処理の高速化を図ることができる。

【0022】また、請求項23記載の発明では、さらに、パターン記憶手段に記憶されているユニットパターンと対応付けされて、条件を含むルールが記憶されるルール記憶手段が設けられており、パターン検索手段は、モーラなどの入力言語情報単位ごとにルール記憶手段に

記憶されているルールの条件を調べることによって、ルールの条件を満たすユニットパターンをパターン記憶手段から検索するようになっている。ユニットパターン検索処理をルール検索とすることによって、ユニットパターンの追加、削除、変更などのメンテナンスが容易となる。

【0023】また、請求項24記載の発明では、ルールの条件には、言語情報単位の特徴を表わす複数の変数の値の範囲が含まれており、パターン検索手段は、ルールの条件に含まれる全ての変数について、入力言語情報単位の値がその範囲内にあるとき当該ルールを採用することによって、ユニットパターンを検索するようになっている。ルールの条件に言語情報単位の特徴を表わす複数の変数の値の範囲を含ませることで、ルール検索をルーチン処理で行なうことができる。また、ルールを外部データ化することが容易となり、ルール自体の保守性を高めることができる。

【0024】また、請求項25記載の発明では、ルールの条件に含まれる変数として、当該言語情報単位の句頭あるいは句末からの位置を用いるか、または、当該言語情報単位を含む句の隣接句との境界種別を用いるか、または、当該言語情報単位を含む句のアクセント情報を用いるか、または、当該言語情報単位の音韻情報を用いるようになっている。ルールの条件に含まれる変数として、言語情報単位の句頭あるいは句末からの位置を用いることによって、句頭から句末へ向かって徐々に下降するなど、句内位置によって連続的に変化するピッチパターンの性質を表現することが可能となる。また、言語情報単位を含む句の隣接句との境界種別を用いることによって、読点位置よりも文末を特に下げることによる文の終始感の演出や、疑問文等の様々なイントネーションスタイルへの対応など、句境界位置でのピッチパターンの特徴を表現することが可能となる。また、言語情報単位を含む句のアクセント情報を用いることによって、頭高型が中高型よりも左に傾いた山になるなどのアクセント型による特徴や、プロミネンス、副次アクセント等のアクセントレベルによる特徴など、ピッチパターンの特徴を表現することが可能となる。また、言語情報単位の音韻情報を用いることによって、特定の音韻で見られる局所的なピッチパターンの特徴や、音韻継続時間長の影響などを表現することが可能となる。

【0025】また、請求項26記載の発明では、1つのルールに複数のユニットパターンが対応付けされており、ルールの条件を満たすユニットパターンが複数検索される場合に、ゆれを与えて1つのユニットパターンを選択するようになっている。これにより、画一的で機械的なピッチパターンを避けることが可能となり、より自然なピッチパターンを得ることができる。

【0026】

【実施例】以下、本発明の実施例を図面に基づいて説明

する。図1は本発明に係るピッチパターン生成装置の一実施例のブロック図である。本実施例のピッチパターン生成装置は、自然音声のピッチパターンから所定の単位で切り出したユニットパターンが記憶されるパターン記憶部1と、所定の入力属性に基づき、該入力属性に適合したユニットパターンを検索するパターン検索部2と、検索により得られたユニットパターンを時間軸上に配置するパターン配置部3と、時間軸上に配置されたユニットパターンを変形するパターン変形部4と、変形されたユニットパターンを接続して一連のフレーズのパターンを生成するパターン接続部5とを有している。

【0027】ここで、パターン記憶部1に記憶されるユニットパターンは、自然音声（すなわち人間の発声）のピッチパターンから図2、図3、図4または図5に示すような仕方で、切り出しがなされる。すなわち、図2では、文発声されたピッチパターンPPのアクセント句AKの上昇部AUの始点から平坦部AFを経て下降部ADの終点までを単位DM₀としてユニットパターンUNPの切り出しを行なうようになっている。また、図3では、文発声されたピッチパターンPPのアクセント句AKの上昇部AUの始点からそれに続く平坦部AFの終点までと、下降部ADの始点からそれに続く平坦部AFの終点までとを、それぞれ単位DM₁、DM₂として、ユニットパターンUNPの切り出しを行なうようになっている。また、図4では、平坦部AFだけを単位DM₀として、ユニットパターンUNPの切り出しを行なうようになっている。また、図5では、文末のイントネーションを表現するような特徴のあるピッチ変化部分を単位DM₀としてユニットパターンUNPの切り出しを行なうようになっている。なお、図2乃至図5において、アクセント句の平坦部AFとは、ピッチの上昇部AUの終点から下降部ADの始点までの山の部分と、その反対の谷の部分との両方を指している。また、図2乃至図5に示したような切り出しは、アクセント核位置などの属性情報との関連においてなされるが、この他にも、図示しないが、特定の範囲のピッチ変化率を示すピッチパターン区間を単位としてユニットパターンを切り出しても良いし、ある特定の範囲の曲率を示すピッチパターン区間を単位としてユニットパターンを切り出しても良い。

【0028】このようにして、パターン記憶部1には、上記いずれか1つの仕方によって文音声から切り出された種々のユニットパターンが記憶されているか、あるいはこれらのユニットパターンを予め作成したコードブック等によってベクトル量子化したユニットパターンがデータベース化して記憶されている。

【0029】また、パターン検索部2は、所定の入力属性に適合したユニットパターンをパターン記憶部1から検索するが、この検索の仕方としては、パターン記憶部1内に記憶されている各ユニットパターンに、これを切り出したときの属性情報を付加しておき、入力属性に最

も類似する属性情報をもつユニットパターンを選ぶという方法が考えられる。さらに、ユニットパターン切り出しの際に、予め作成したコードブックによってユニットパターンがベクトル量子化されてパターン記憶部1に記憶されている場合には、ベクトル量子化されたユニットパターンのコードを出力としたそのときの属性を入力とするような教師データによって学習させたニューラルネットワークに、所定の入力属性を入力してユニットパターンを検索する方法も考えられる。

【0030】また、パターン検索部2において用いられる所定の入力属性とは、規則合成器に与えられるような入力情報のうち、ピッチパターンに影響を与えられられる情報であり、代表的なものとしてアクセント型などがあるが、この他に本実施例では以下の情報も入力属性に加える。すなわち、従来から、ピッチが変化する前後の近傍の音韻の種類がピッチの局所的変化をもたらすことが知られており、この情報をパターン検索に用いることが有効と考えられ、これを入力属性に加える。また、従来から、プロミネンスが付加された箇所はアクセントの高さなどに変化を生ずることが知られているため、この情報をパターン検索に用いることが有効と考えられ、これを入力属性に加える。また、図5に示したようなパターンを選択する場合には、フレーズのイントネーションスタイルの情報が必要であり、この情報をパターン検索に用いることが有効と考えられ、これを入力属性に加える。

【0031】また、パターン配置部3は、アクセント句境界の時間軸上の位置を入力情報として、あるいは、アクセント核の時間軸上の位置を入力情報として、あるいは、フレーズ末の音韻、音節、モーラなどの境界の時間軸上の位置を入力情報として、あるいは、ピッチ変化部分近傍の音韻環境情報を入力情報としてユニットパターン時間軸上の位置を決定するようになっている。

【0032】また、パターン変形部4は、各ユニットパターンの終点を次のユニットパターンの始点に一致させるように、使用されたユニットパターンの種類等に応じ、アクセント句内の上昇部AU、平坦部AF、下降部ADのいずれか1つあるいは2つあるいは全てを、時間軸方向および/またはピッチ軸方向に伸縮するような変形を行なうようになっている。

【0033】また、パターン接続部5は、使用されたユニットパターンの種類等に応じ、ユニットパターン間を直線で、あるいは3次曲線で補間して接続するようになっている。

【0034】次に、このような構成のピッチパターン生成装置の動作について説明する。先づ、人間の発声からユニットパターンを切り出して、これを予めパターン記憶部1に記憶する。ユニットパターンの切り出しは、前述したように、例えば図2、図3、図4または図5に示したような方法によって行なうことができる。

【0035】すなわち、図2のような方法が用いられる場合、アクセント句の上昇部の始点から平坦部を経て下降部の終点までを単位としてユニットパターンを切り出すことができる。なお、これに類するものとして、アクセント句をそのまま単位として切り出すことも考えられ、この場合にも図2と同様に、アクセント句のパターンをほぼそのまま記憶しておくことができるが、図2に示した仕方ではこれに比べて、同アクセント型の異なるモーラ数の句を同列に扱えるという利点がある。

【0036】また、図3のような方法が用いられる場合、アクセント句の上昇部の始点からそれに続く平坦部の終点までと、下降部の始点からそれに続く平坦部の終点までをそれぞれ単位としてユニットパターンを切り出すことができる。なお、これに類するものとして上昇部、下降部のみを単位として切り出すことも考えられ、この場合にもユニットパターンの始点の位置を決定するだけで、ユニットパターンをほぼ配置することができるが、図3に示した仕方はこれに比べて、パターン間接続の際の補間によって生じる自然性劣化を低減できるという利点がある。

【0037】また、図4のような仕方が用いられる場合、平坦部だけを単位としてユニットパターンを切り出すことができる。文献「“基本周期のゆらぎの性質とそのモデルに関する検討”、小室修、粕谷英樹、日本音響学会誌、Vol. 47, No. 12, pp. 928-934, 1991」には、定常母音部分での遅いゆらぎが自然性に影響を与えると報告されており、図4の方法によって切り出しを行えば、このような聴覚上重要なゆらぎを保存することができる。

【0038】また、図5のような仕方が用いられる場合、文末のイントネーションを表現するような、特徴あるピッチ変化部分を単位としてユニットパターンを切り出すことができる。なお、図5では、疑問を表わす尻上がりのパターンの例が示されている。図5の方法によれば文末のイントネーションのパターンをそのまま記憶できるため、規則では表現が困難な微妙なニュアンスの違いも表現が可能になる。

【0039】図2乃至図5の例では、アクセント核位置などの属性情報との関連において切り出されたユニットパターンを使用することを前提としたが、この他に、特定の範囲のピッチ変化率や曲率をもつパターンを自動的に切り出したユニットパターンを使用する方法も考えられる。前者はピッチが変化する部分をパターンとして記憶することによってその外形を保存できることや、定常部をパターンとして記憶することによってゆらぎを保存できるという利点がある。また、後者では曲線状の部分をパターンとして記憶することによって、その他の部分を直線で補間できるため、処理量を低減することができる。図2乃至図5あるいはその他の方法によって切り出されたユニットパターンは、これをそのままパターン記

憶部 1 に記憶しても良いし、あるいは、ベクトル量子化などを行なってコード化してパターン記憶部 1 に記憶しても良い。

【0040】このようにして、パターン記憶部 1 にユニットパターンを記憶した後、実際のピッチパターン生成処理を開始させることができる。このピッチパターン生成処理では、先づ、パターン検索部 2 に所定の入力属性を与え、この入力属性に適するユニットパターンをパターン記憶部 1 から検索させる。ここで、本実施例では、入力属性には、代表的なアクセント型などの他に、ピッチが変化する前後の近傍の音韻の種類の情報、プロミネンスの情報、フレーズのイントネーションスタイルの情報のうちの少なくとも 1 つの情報が付加されている。このような入力属性の種々の情報のうち、ピッチ変化部分近傍の複数モーラにわたる音韻環境を用いユニットパターンを検索する場合には、音韻環境によるパターンの局所的変化の再現性が向上し、より緻密なパターン生成が可能になる。また、プロミネンスの情報を用いユニットパターンを検索する場合には、従来単調になりがちだった合成音声の抑揚にめりはりがつき、了解性、自然性を向上させることができる。また、イントネーションスタイルを用いユニットパターンを検索する場合には、規則による生成では表現が困難なイントネーションを正確に再現することが可能となる。また、ニューラルネットワークによってユニットパターンを検索することによって、学習した人間の発声パターンに近いパターンを再現することができ、自然性を向上させることができる。

【0041】このようにして、最適なユニットパターンが検索された後、パターン配置部 3 では、主にアクセント核やアクセント句境界の時間軸上の位置などから、検索されたユニットパターンの生成ピッチパターンにおける時間軸上の位置、より詳しくはユニットパターンの始点の位置を決定し、ユニットパターンを配置する。より詳しくは、パターン配置部 3 は、ユニットパターンの始点の位置を記号レベルでのピッチレベルが変化する音韻境界位置に、その部分の音韻環境によって変化するずれを考慮して決定する。ここで、ピッチレベルが変化する位置とは、アクセント句に関しては句境界や核位置などであるが、図 5 に示したようなユニットパターンに関しては、例えば疑問を表わす尻上がりのパターンなどは最終母音開始点を基準として位置が決まることが文献

「“日本語疑問文の知覚”，三浦一郎，杉藤美代子，日本音響学会秋季研究発表会講演論文集，1991，1，2-6-9，p255-256」などによって知られているので、フレーズ末の音韻，音節，モーラなどの境界位置を用いる。図 6 は図 3 に示したようなユニットパターンを用いたときのパターン配置処理の様子を示す図であり、図 6 において、下部には、音韻境界の位置と記号レベルでのピッチレベルの高低を示し、上部にユニットパターンを配置した様子を示している。図 6 において、点線 L 1 は上昇部ある

いは下降部と平坦部との境界を表わしており、この場合、ユニットパターンの始点の位置は太線 L 2 のように決定される。

【0042】このようにして、ユニットパターンの時間軸上の位置を決定した後、パターン変形部 4 では、生成ピッチパターンに自然に埋めこまれるように、ユニットパターンを変形する処理を行なう。すなわち、パターン変形部 4 では、先づ、時間軸上の変形を行なう。図 7 は図 6 のように配置のなされたユニットパターンに対して変形を施す様子を示す図である。すなわち、図 6 では、図 3 に示したようなユニットパターンを使用しているので、図 7 では各ユニットパターンの始点を前のユニットパターンの終点に一致させるように、時間軸方向の伸縮を行なう。このとき、時間軸方向の伸縮は平坦部のみで行なうことが、発声機構の生理的制約からも妥当と考えられる。なお、ここで、伸縮の具体的方法は特に限定するものではなく、また、伸縮には、データの打ち切りや繰り返し、外挿なども含まれる。

【0043】このようにして、時間軸方向の伸縮を行なった後、ピッチ軸上の変形を行なう。図 8 は図 7 のように時間軸方向の変形がなされたピッチパターンにピッチ軸上の変形を施す様子を示す図である。すなわち、この例では、各ユニットパターンの始点を前のユニットパターンの終点に一致させるように、ピッチ軸方向の伸縮がなされる。このとき、上昇部，下降部は前後のつながりによって変化することは妥当であるので、ピッチ軸方向の伸縮を、上昇部および下降部のみで行なうのが良く、これによって、平坦部の概形やゆらぎを保存することができる。このような方法によってユニットパターンの平坦部はその絶対的高さも保存される。このため、プロミネンスの影響やフレーズ中の位置などによる影響も含めて、ピッチパターンをユニットパターンによって再現することができる。

【0044】上述の例では、図 3 に示したようなユニットパターンを使用したため、パターン間の接続に際して補間処理は特に必要としなかったが、図 3 に示したようなユニットパターン以外のユニットパターンを使用する場合には、パターン接続部 5 において何らかの補間処理が必要となる。このとき、上昇部と下降部のみをユニットパターンとする場合にはその間を直線によって、また、図 4 のように平坦部をユニットパターンとする場合にはその間を 3 次曲線によって補間することで、より人間のピッチパターンに近いパターンを生成することができる。

【0045】ところで、上述したようなユニット接続型のピッチパターン生成装置において、どのようなユニットパターンをパターン記憶部に記憶するかという問題が性能を大きく左右する。そこで本発明では、さらに、パターン記憶部に記憶されるユニットパターンとして、特定の言語的状況において特徴的に現われるピッチパター

ンを用い、これにより、効率的で効果的なピッチパターン生成を実現するようにしている。以下、このようなユニットパターンについて詳述する。

【0046】連続音声には呼吸段落や文などの複数のフレーズがあるが、通常の発声では、フレーズ内位置が末尾に近づくにつれて低いピッチパターンとなる。従って、フレーズ内位置の異なる複数のパターンを用意することにより、声立て成分の分離・合成などの複雑な処理をせずに、上記のようなフレーズ感を出すことが可能となる。また、文末では文の長さにかかわらず、比較的低い専用パターンを使用することによって、文の終始感を出すことができる。このように、ユニットパターンの1つとして、呼吸段落や文など、一連のフレーズにおける特定の位置において特徴的に現われるピッチパターンを用いることにより、フレーズ感、文の終始感などを表現することができる。

【0047】また、アクセント句のパターン概形は、アクセント核の有無、位置などによって異なる。無核のアクセント句パターンは有核のそれに比べてなだらかであり、頭高型のもはや急峻に上昇してなだらかに下降する。従って、アクセント型の異なる複数のパターンを用意し使い分けることで、このような違いを表現できる。このように、ユニットパターンの1つとして、頭高型、平板型など特定のアクセント型において特徴的に現われるピッチパターンを用いることにより、より自然なピッチパターンを生成することができる。

【0048】また、プロミネンスが付与されている部分は、比較的高い専用パターンを使用することで、フレーズ頭でも不自然なピッチ上昇のない自然なプロミネンス感を出すことができる。また、従来困難であった副次アクセントの表現も、アクセントレベルの異なるパターンを用意することで可能となる。このように、ユニットパターンの1つとして、プロミネンスが付与されている部分、副次アクセント核が現われる部分などの特定のアクセントレベルにおいて特徴的に現われるピッチパターンを用いることにより、プロミネンス、副次アクセントの表現が可能となり、自然性を向上させることができる。

【0049】また、ユニットパターンの1つとして、間投詞や終助詞、擬声語、擬態語、掛け声などのその語に独特のピッチパターンで発声される語を用いるか、あるいはフレーズのピッチパターンをそのまま用いることもでき、この場合には、その語に独特のピッチパターンで発声される語を忠実に再現できて、自然性を向上させることができる。

【0050】また、数字を棒読みする場合や和歌などの韻文を朗読する場合は、一定のピッチパターンが周期的に現われる。従って、ユニットパターンの1つとして、数字列、韻文などに周期的に現われる特徴的なピッチパターンを用いることにより、上記のような独特の節回しの発声にも対応することができる。

【0051】また、会話では「えー」や「うん」など、語そのものにはあまり意味はなく、主に韻律の違いによって意図を表現する場合が多々ある。従って、ユニットパターンの1つとして、語や音韻から独立して主に韻律によって特定の意図を表わす特徴的なピッチパターンを用いることにより、効率的に自然な会話出力ができる。

【0052】図9は、入力言語情報からこれらのユニットパターンを検索してピッチパターンを合成する処理の具体例を示す図である。図9の例では、パターン記憶部1に、【P01】文頭平板型、【P03】文末平板型、【P41】プロミネンス、【P42】副次アクセント、【P51】ですか？、【P61】数字、【P71】意外、の7種類のユニットパターンが予め記憶されている場合が示されており、入力言語情報として、「え？（意外）／920917ですか？ もう一度／メニュー番号を（プロミネンス）／入力してください」が入力すると、パターン検索部2では、パターン記憶部1からこの入力言語情報に対応したユニットパターン列を検索する。この結果、「【P71】え？【P61】92【P61】09【P61】17【P51】ですか？【P01】もう一度【P41】メニュー番号を【P03】入力して【P42】ください」のユニットパターン列を検索することができる。次いで、このように検索されたユニットパターン列をパターン配置部3、パターン変形部4、パターン接続部5において、配置、変形、接続することによって、図9の下部に示すような自然なピッチパターンを生成することができる。

【0053】このように、パターン記憶部1に記憶されるユニットパターンとして、特定の言語状況において特徴的に現われるピッチパターンを用いることにより、入力言語情報からヒューリスティックにユニットパターンを検索することができ、このユニットパターンを配置、変形、接続することによって、比較的簡単な処理で効率的かつ効果的に、自然なピッチパターンを生成することができる。

【0054】なお、ユニットパターンとして自然音声のピッチパターンをそのまま記憶する場合には、生成規則やモデルパラメータとして記憶するよりも大量のデータが必要となり、メモリ容量の増加や処理量の増加によるリアルタイム処理の困難が予想される。このため、それを回避するような何らかの工夫が必要となる。

【0055】本願の発明者は、メモリ容量の増加や処理量の増加によってリアルタイム処理が困難になるのを防止するため、自然音声のピッチパターンから適当な単位で切り出したユニットパターンをそのままの形でパターン記憶部1に記憶するのではなく、ユニットパターンを折れ線近似してパターン記憶部1に記憶することをさらに案出した。

【0056】図10には、本来曲線的であるユニットパターンを折れ線近似した一例が示されている。図10の

例では、1つのユニットパターンが、7つの折れ線に近似され、各折れ線ごとの区間（セグメント）S1及至S7に区分されている。ここで、ユニットパターンは、始点のピッチと、それぞれの区間の区間長（セグメント長）と、それぞれの区間の折れ線（線分）の傾きを表すステップ値とによって表現することができ、ユニットパターンをこの表現形式でパターン記憶部1に記憶することができる。なお、ステップ値とは、ピッチ周期が1サンプル変化するのに要する時間、すなわちピッチ更新周期であり、折れ線の傾きが図11(a)のように右上がりのときには正(+)の値をとり、また、図11(b)のように右下がりのときには負(-)の値をとるものとする。

【0057】このように、本来曲線的であるユニットパターンを折れ線で近似し、これを始点のピッチおよびそれぞれの区間の区間長とステップ値とにより表現してパターン記憶部1に記憶することにより、ユニットパターンをそのままの形で記憶する場合に比べて、パターン記憶部1に記憶されるデータ量を大幅に減少することができる。さらに、ユニットパターンのピッチ方向のシフトを始点ピッチの増減のみで行なうことができ、また、時間軸方向の伸縮を区間長の増減によって行なうことができ、パターンの変形も簡単に行なうことができる。

【0058】但し、上記のような表現を用いる場合には、パターン検索部2において、パターン記憶部1から所定のユニットパターンの上記表現形式のデータを検索し読み出したときに、これをユニットパターンの形に（すなわち折れ線の形に）再生（生成）する必要があるが、このユニットパターン生成処理は、図12に示すような簡単な構成によって実現できる。すなわち、図12に示す回路は、ある1つの区間における折れ線を再生（生成）するのに、その区間のステップ値が入力するようになっており、入力したステップ値の絶対値を求める（すなわち、ステップ値が負のときに符号を反転させる）インバータ11と、ステップ値の絶対値が設定され、所定時間間隔のサンプルクロックを計数し、サンプルクロックの計数値がステップ値の絶対値が一致したときに、“1”を出力するステップカウンタ12と、入力したステップ値の正負を判定するコンパレータ13と、コンパレータ13において正(+)と判別され、ステップカウンタ12から“1”が出力されたときに“1”を出力するAND回路14と、コンパレータ13において負(-)と判別され、ステップカウンタ12から“1”が出力されたときに“1”を出力するAND回路15と、AND回路14からの出力“1”を正(+)方向にカウントし、また、AND回路15からの出力“1”を負(-)方向にカウントし、そのカウント値をピッチとして出力するピッチカウンタ16とから構成されており、この回路によって、ステップ値（ピッチ更新周期）が正の場合、ピッチ周期を単位長増加させ、ステップ値

が負の場合、ピッチ周期を単位長減少させる処理を行なうことができる。換言すれば、ピッチカウンタ16の初期値に始点のピッチを設定し、1つの区間において時間軸方向に区間長（セグメント長）となるまで上記回路のピッチカウンタ16からピッチを出力させ、これを各区間について順次に行なうことにより、折れ線近似のユニットパターンを再生（生成）することができる。このようにして、ユニットパターンを再生することができ、また、実際のピッチパターンを生成することができる。

【0059】また、パターン配置部3では、パターン検索部2で検索され生成された各ユニットパターンを所定位置に配置し、パターン変形部4では、自然なピッチパターンを得るために、パターン配置部3により配置された各ユニットパターンを変形する処理を行なうが、このときに、隣接ユニットパターン間の接続やプロミネンス付与などのために、傾きを調整する必要がある。本発明では、前述のようなデータ表現を採用しているため、この調整処理は、例えば、過渡部（図10を参照）の全区間のステップ値に一定のバイアスを加算器（図示せず）により加えることで、容易に実現できる。

【0060】図13には、1つのユニットパターンに対するこのような調整処理例が示されている。図13を参照すると、1つのユニットパターンの各区間のステップ値のそれぞれに一定のバイアス値を加えて、ステップ値を変えることによって（各折れ線の傾きを変えることによって）、ユニットパターンにピッチ方向の滑らかな変形を施すことができる。例えば、図13の例のように、ユニットパターンの各区間のステップ値（変形前） ST_i に負のバイアス値BIASを加えて ST_i' とすることによって、変形前のユニットパターン UNP_i を UNP_i' のように変形することができる。すなわち、終点については変動させず、始点を所定量（バイアス量）変動させることができる。このようにして、ユニットパターン間の接続やプロミネンス付与などで必要となるパターン変形処理が加算器だけの非常に簡単な構成で実現できる。

【0061】なお、このような仕方では自然な傾きの調整を行なうには、右上がりの傾きから水平、右下がりの傾きへと連続的に変化させる必要があるが、これは、ステップ値の最大値（正の最大絶対値）と最小値（負の最大絶対値）とを連続的に変化させるようにすることで可能となり、このような処理は、ステップ値（ピッチ更新周期）を“2”の補数表現によって扱い、バイアスを加えた結果のステップ値の絶対値が最大値を越えた場合に、これを反対符号の最大絶対値とすることによって、容易に実現することができる。

【0062】このように、パターン変形部4は、1つのユニットパターンの各区間のユニット値に一定のバイアスを加えることによって、このユニットパターンをピッチ方向に滑らかに変形することができるが、パターン変形部4では、隣接ユニットパターン間の接続やプロミネ

ンス付与などで必要となるバイアスをリアルタイムに計算しなければならない。この計算には通常乗除算を要するために処理が複雑となり、ピッチパターン生成処理全体をリアルタイムに行なう上でボトルネックとなる場合がある。この問題を回避するには、ピッチ変化量（例えば原パターンと生成パターンの始点ピッチの差、あるいは生成パターンの始点ピッチそのもの）とバイアスとの対応関係を予め計算した図14に示すようなバイアステーブルをユニットパターン毎に設け、ある1つのユニットパターンをピッチ方向に所定のピッチ変化量で変形するときには、このユニットパターンのバイアステーブルから上記ピッチ変化量に対応したバイアスを読み出して用いられれば良い。これによって、バイアスをその都度計算する必要をなくし、リアルタイム処理に支障を生じさずに済む。具体的には、図15に示すように、各パターンの終点についてはこれを変動させず、各パターンの始点が前のパターンの終点となるように始点のピッチ変化量を割り出し、これに対応したバイアスを図14に示すようなバイアステーブルから読み出して用い、原パターンを変形して生成パターンとすることで、前のパターンとの接続を容易に行なうことができる。

【0063】また、本発明の変形例として、パターン検索部2におけるユニットパターンの追加、削除、変更等のメンテナンスを容易にするため、図16に示すように、入力言語情報単位ごとにユニットパターンを検索するための複数のルールを記憶しているルール記憶部7をさらに設けることもできる。

【0064】この場合、ルール記憶部7に記憶されるルールとしては、複数のアイテムとその値の範囲とからなる条件と、ユニット番号とにより構成されたものを用いることができる。ここで、アイテムは、モーラごとの各種属性であって、具体的には、句頭からの位置、隣接句との境界種別、アクセント型、アクセントレベル、音韻などが有効と考えられる。

【0065】次に図16に示すような構成のピッチパターン生成装置におけるパターン検索処理の具体例を図17乃至図19を用いて説明する。なお、図17はパターン記憶部1に記憶されている各種ユニットパターンを示す図であり、図17の例では、4種類のユニットパターンがユニット番号“1”、“2”、“3”、“4”を付されて記憶されている。また、図18はルール記憶部7に記憶されているルールの構成例を示す図であり、図18の例では、ルールは、アイテムとして、アクセント句頭からの位置Pos、声立て句頭からの位置Mora、アクセント型Acc、アクセント核の位置Coreの4つの属性を用いており、4種類のユニット番号“1”、“2”、“3”、“4”のそれぞれに、4つの属性の値の範囲が割り当てられたものとなっている。

【0066】図16において、パターン検索部2では、図18のルール記憶部7を参照して、パターン記憶部1

に記憶されている4種類のユニットパターンから、入力言語情報に応じたユニットパターンを検索する。具体的には、入力言語情報単位をモーラとし、処理がモーラ単位に行なわれるとするとし、図19に示すような入力モーラ列が入力する場合には、先ず、各モーラについて、すべてのアイテムの値を求める。ある1つのモーラに着目すると、パターン検索部2は、ルール記憶部7を参照し、ルールの条件に含まれる全てのアイテムについて、その範囲に、このモーラの各アイテムの値が入っているか否かを調べる。そして、ある1つのユニット番号のアイテムの範囲にこのモーラの各アイテムの値が入っていると判断されたときには、このモーラは、ルールの条件を満たすものとし、パターン検索部2は、そのルールのユニット番号に対応したユニットパターンをパターン記憶部1から検索することができる。また、条件を満たすルールが1つもないときには、そのモーラ位置に配置されるユニットパターンはないものとして、次のモーラについての処理に進む。

【0067】例えば、図19の入力モーラ列において、「メロン」の「ロ」というモーラは、各アイテム値が“0”、“4”、“1”、“0”であって、ルール記憶部7に記憶されているルールのうち、ユニット番号“4”のルールの条件を満たすので、パターン検索部4は、パターン記憶部1からユニット番号“4”のユニットパターンを検索することができる。

【0068】このように、ルールの条件に言語情報単位の特徴を表わす複数の変数の値の範囲を含ませることで、ルール検索をルーチン処理で行なうことができる。また、ルールを外部データ化することが容易となり、ルール自体の保守性を高めることができる。

【0069】また、ルールの条件に含まれる変数として、言語情報単位の句頭あるいは句末からの位置を用いることによって、句頭から句末へ向かって徐々に下降するなど、句内位置によって連続的に変化するピッチパターンの性質を表現することが可能となる。また、言語情報単位を含む句の隣接句との境界種別を用いることによって、読点位置よりも文末を特に下げることによる文の終始感の演出や、疑問文等の様々なイントネーションスタイルへの対応など、句境界位置でのピッチパターンの特徴を表現することが可能となる。また、言語情報単位を含む句のアクセント情報を用いることによって、頭高型が中高型よりも左に傾いた山になるなどのアクセント型による特徴や、プロミネンス、副次アクセント等のアクセントレベルによる特徴など、ピッチパターンの特徴を表現することが可能となる。また、言語情報単位の音韻情報を用いることによって、特定の音韻で見られる局所的なピッチパターンの特徴や、音韻継続時間長の影響などを表現することが可能となる。

【0070】なお、上記例では、ルールとユニットパターンは1対1対応のものとなっているが、1つのルール

10

20

30

40

50

が複数のユニット番号をもつこともある。この場合には、その中から何らかの“ゆれ”を与えて1つのユニットパターンを選択することもできる。これによって、同じ入力からでも全く同じではないピッチパターンを生成することができ、画一的で機械的なピッチパターンを避けることが可能となり、より自然なピッチパターンを得ることができる。

【0071】

【発明の効果】以上に説明したように、請求項1記載の発明によれば、人間の発声のピッチパターンから適当な単位で切り出したユニットパターンをパターン記憶手段に記憶し、パターン検索手段では、所定の入力属性に基づき、該入力属性に適合したユニットパターンをパターン記憶手段から検索し、パターン配置手段では、検索により得られたユニットパターンを時間軸上に配置し、パターン変形手段では、時間軸上に配置されたユニットパターンを変形して、ピッチパターンを生成するようにしており、人間の発声したピッチパターンをそのまま使用しているため、規則による生成では困難な自然な変化やゆらぎ、ニュアンスなどを表現可能なピッチパターンを生成することができる。

【0072】また、請求項2記載の発明において、アクセント句の上昇部と下降部とをそれぞれ1つ含むものを1つの単位として切り出したものをユニットパターンとして使用する場合には、ピッチパターンの重要な単位であるアクセント句のパターンをほぼ保存したピッチパターンを生成することができる。また、アクセント句の上昇部または下降部を1つ含むものをユニットパターンとして使用する場合には、簡単な接続処理によってピッチパターンを生成することができる。また、アクセント句の平坦部を含むユニットパターンを使用する場合には、定常部のゆらぎ等を保存でき、自然なピッチパターンを生成することができる。また、フレーズ末の特徴的ピッチ変化をユニットパターンとして使用する場合には、規則による生成では困難な多彩なイントネーションを表現することができる。また、特定の範囲のピッチ変化率を示すパターンをユニットパターンとして使用する場合には、変化部分の概形や定常部分のゆらぎ等を保存することができ、規則による生成では困難であった多彩で自然なピッチパターンを生成することができる。また、特定の範囲のピッチ曲率を示すパターンをユニットパターンとして使用する場合には、その他の部分を直線で補間できるため、処理量を低減することができる。

【0073】また、請求項3記載の発明において、ベクトル量子化されたユニットパターンをパターン記憶手段に記憶し、これを使用するときには、パターン記憶手段の容量を大幅に削減することができ、装置の低コスト化、小型化、処理の高速化を図ることができる。

【0074】また、請求項4記載の発明によれば、ピッチ変化部分近傍の複数モーラにわたる音韻環境からユ

ットパターンを検索することによって、音韻環境によるパターンの局所的变化の再現性が向上し、より緻密なピッチパターンを生成することができる。また、プロミネンスの情報からユニットパターンを検索する場合には、従来単調になりがちだった合成音声の抑揚にめりはりがつき、了解性、自然性を向上させることができる。また、イントネーションスタイルからユニットパターンを検索する場合には、規則による生成では表現が困難なイントネーションを正確に再現することができる。

【0075】また、請求項5記載の発明によれば、ニューラルネットワークによってユニットパターンを検索することによって、学習した人間の発声パターンに近いパターンを再現することができ、自然性を向上させることができる。

【0076】また、請求項6記載の発明において、アクセント句境界の位置を用いてユニットパターンを配置する場合には、アクセント句の上昇開始点をほぼ再現でき、了解性を向上させることができる。また、アクセント核の位置を用いてユニットパターンを配置する場合には、アクセント句の下降開始点をほぼ再現でき、了解性を向上させることができる。また、フレーズ末の音韻、音節、モーラなどの境界位置を用いてユニットパターンを配置する場合には、イントネーションを効果的に表現でき、自然性を向上させることができる。また、ピッチ変化部分近傍音韻環境情報を用いてユニットパターンを配置することにより、ピッチパターンの時間構造をより緻密に再現でき、自然性を向上させることができる。

【0077】また、請求項7記載の発明によれば、アクセント句の平坦部を時間軸方向に伸縮するようにしており、ピッチ変化部については時間軸方向に変形しないことにより、ピッチ変化部の時間構造を保存することができ、生理的制約から外れないパターンが保証されるため、安定した自然なピッチパターンを生成することができる。また、アクセント句の上昇部と下降部をピッチ軸方向の変形をするようにしており平坦部についてはピッチ軸方向に変形しないことにより、ピッチ定常部部の高さやゆらぎ等を保存することができ、プロミネンスなども含めたパターン表現が可能となり自然性を向上させることができる。

【0078】また、請求項8、9記載の発明によれば、ユニットパターン間を直線で補間することによって、処理量を低減でき、高速化や低コスト化を図ることができる。また、ユニットパターン間を3次曲線で補間することによって、人間のピッチ変化に似た概形を表現でき、自然なピッチパターンを生成することができる。

【0079】また、請求項10記載の発明によれば、パターン記憶部に記憶されるユニットパターンとして、特定の言語的状況において特徴的に現われるピッチパターンを用いるようになっているので、入力言語情報からヒューリスティックにパターンを検索することができ、効

率的で効果的にピッチパターンを生成することができる。

【0080】また、請求項11記載の発明によれば、上記ユニットパターンの1つとして、呼気段落や文など、一連のフレーズにおける特定の位置において特徴的に現われるピッチパターンを用いることにより、声立て成分の分離・合成などの複雑な処理をせずに、フレーズ感、文の終始感など表現でき、処理の簡略化を図ることができる。

【0081】また、請求項12記載の発明によれば、上記ユニットパターンの1つとして、頭高型、平板型など特定のアクセント型において特徴的に現われるピッチパターンを用いることにより、アクセント核の有無、位置などによって異なるアクセント句のパターン概形を、アクセント型の異なる複数のパターンを用意し使い分けることで表現でき、自然性を向上させることができる。

【0082】また、請求項13記載の発明によれば、上記ユニットパターンの1つとして、プロミネンスが付与されている部分、副次アクセント核が現われる部分などの特定のアクセントレベルにおいて特徴的に現われるピッチパターンを用いることにより、プロミネンス、副次アクセントの表現も可能となり、自然性を向上させることができる。

【0083】また、請求項14記載の発明によれば、上記ユニットパターンの1つとして、間投詞や終助詞、擬声語、擬態語、掛け声などのその語に独特のピッチパターンで発声される語を用いるか、あるいはフレーズのピッチパターンをそのまま用いることにより、その語に独特のピッチパターンで発声される語を忠実に再現でき、自然性を向上させることができる。

【0084】また、請求項15記載の発明によれば、上記ユニットパターンの1つとして、数字列、韻文などに周期的に現われる特徴的ピッチパターンを用いることにより、数字を棒読みする場合や韻文を朗読する場合などの独特の節回しの発声にも対応することができ、装置の汎用性を高めることができる。

【0085】また、請求項16記載の発明によれば、上記ユニットパターンの1つとして、語や音韻から独立して主に韻律によって特定の意図を表わす特徴的ピッチパターンを用いることにより、会話独特の「えー」や「うん」など、主に韻律の違いによって意図を表現する場合にも、効率的に自然な会話出力が可能となる。

【0086】また、請求項17、請求項18記載の発明によれば、パターン記憶手段には、折れ線近似されたユニットパターンが記憶されるようになっており、この際、折れ線近似されたユニットパターンは、始点のピッチおよび各折れ線のそれぞれの区間と折れ線の傾きとの表現データによって表現されてパターン記憶手段に記憶されることにより、ユニットパターンのデータ量を減らし、メモリ量と処理量を軽減でき、低コスト化、高速化

を図ることができる。さらに、ユニットパターンのピッチ方向のシフトは始点ピッチの増減のみで、時間方向の伸縮は区間長の増減によって、それぞれ簡単に実現可能で、ユニットパターンの変形が簡単な処理によって実現できるため、高速化、高機能化を図ることができる。

【0087】また、請求項19記載の発明によれば、上記各折れ線の傾きは、ピッチ周期が単位長だけ変化するのに要する時間としてのステップ値によって表現されており、この場合、パターン検索手段は、パターン記憶手段に記憶されているユニットパターンの表現データを検索し、該表現データに基づきユニットパターンを再生するときに、ステップ値が正の値の場合、ピッチ周期を単位長増加させ、ステップ値が負の値の場合、ピッチ周期を単位長減少させる処理を行なうようになっているので、カウンタ、比較器などによる非常に簡単な構成で折れ線近似のユニットパターンを再生し、実際のピッチパターンを得ることができる。

【0088】また、請求項20記載の発明によれば、パターン変形手段は、各折れ線の傾きを表わすステップ値に一定のバイアスを加えることによって、ユニットパターンにピッチ方向の滑らかな変形を施すので、ユニットパターン間の接続やプロミネンス付与などで必要となるパターン変形処理が加算器だけの非常に簡単な構成で実現できる。

【0089】また、請求項21記載の発明によれば、パターン変形手段は、ステップ値を“2”の補数表現によって扱い、バイアスを加えた結果のステップ値の絶対値が最大値を越えた場合に、これを反対符号の最大絶対値とするので、右上がりのパターンから右下がりのパターンへ、またはその逆へ、連続的にかつ無理なく、パターン変形する処理を簡単に実現することができる。

【0090】また、請求項22記載の発明によれば、ピッチ変化量とバイアスとの対応関係が予め求められている場合に、パターン変形手段は、該対応関係を用いてユニットパターンのピッチ方向の変形を行なうので、ユニットパターン間の接続やプロミネンス付与などで必要となるバイアスを、処理量の多い乗除算を必要とすることなく、極めて容易に求めることができ、処理の高速化を図ることができる。

【0091】また、請求項23記載の発明によれば、さらに、パターン記憶手段に記憶されているユニットパターンと対応付けされて、条件を含むルールが記憶されるルール記憶手段が設けられており、パターン検索手段は、モーラなどの入力言語情報単位ごとにルール記憶手段に記憶されているルールの条件を調べることによって、ルールの条件を満たすユニットパターンをパターン記憶手段から検索するようになっており、ユニットパターン検索処理をルール検索とすることによって、ユニットパターンの追加、削除、変更などのメンテナンスが容易となる。

【0092】また、請求項24記載の発明によれば、ルール
の条件には、言語情報単位の特徴を表わす複数の変
数の値の範囲が含まれており、パターン検索手段は、ル
ールの条件に含まれる全ての変数について、入力言語情
報単位の値がその範囲内にあるとき当該ルールを採用す
ることによって、ユニットパターンを検索するようにな
っており、ルールの条件に言語情報単位の特徴を表わす
複数の変数の値の範囲を含ませることで、ルール検索を
ルーチン処理で行なうことができる。また、ルールを外
部データ化することが容易となり、ルール自体の保守性

【0093】また、請求項25記載の発明によれば、ル
ールの条件に含まれる変数として、当該言語情報単位の
句頭あるいは句末からの位置を用いるか、または、当該
言語情報単位を含む句の隣接句との境界種別を用いる
か、または、当該言語情報単位を含む句のアクセント情
報を用いるか、または、当該言語情報単位の音韻情報を
用いるようになっており、ルールの条件に含まれる変数
として、言語情報単位の句頭あるいは句末からの位置を
用いることによって、句頭から句末へ向かって徐々に下
降するなど、句内位置によって連続的に変化するピッチ
パターンの性質を表現することが可能となる。また、言
語情報単位を含む句の隣接句との境界種別を用いること
によって、読点位置よりも文末を特に下げることによる
文の終始感の演出や、疑問文等の様々なイントネーショ
ンスタイルへの対応など、句境界位置でのピッチパター
ンの特徴を表現することが可能となる。また、言語情報
単位を含む句のアクセント情報を用いることによって、
頭高型が中高型よりも左に傾いた山になるなどのアクセ
ント型による特徴や、プロミネンス、副次アクセント等
のアクセントレベルによる特徴など、ピッチパターンの
特徴を表現することが可能となる。また、言語情報単位
の音韻情報を用いることによって、特定の音韻で見られ
る局所的なピッチパターンの特徴や、音韻継続時間長の
影響などを表現することが可能となる。

【0094】また、請求項26記載の発明によれば、1
つのルールに複数のユニットパターンが対応付けされて
おり、ルールの条件を満たすユニットパターンが複数検
索される場合に、ゆれを与えて1つのユニットパターン
を選択するので、画一的で機械的なピッチパターンを避
けることが可能となり、より自然なピッチパターンを得
ることができる。

【図面の簡単な説明】

【図1】本発明に係るピッチパターン生成装置の一実施

例のブロックである。

【図2】ユニットパターンの切り出しの一例を説明する
ための図である。

【図3】ユニットパターンの切り出しの一例を説明する
ための図である。

【図4】ユニットパターンの切り出しの一例を説明する
ための図である。

【図5】ユニットパターンの切り出しの一例を説明する
ための図である。

【図6】ユニットパターンの時間軸上の位置を決定する
様子を説明するための図である。

【図7】ユニットパターンの時間軸上の変形の様子を説
明するための図である。

【図8】ユニットパターンのピッチ軸上の変形の様子を
説明するための図である。

【図9】入力言語情報からユニットパターンを検索して
ピッチパターンを合成する処理の具体例を示す図であ
る。

【図10】ユニットパターンを折れ線近似した一例を示
す図である。

【図11】(a)、(b)はステップ値を説明するための
図である。

【図12】ユニットパターン生成回路の一例を示す図で
ある。

【図13】ユニットパターンの傾きを調整する処理を説
明するための図である。

【図14】パイアステープルの一例を示す図である。

【図15】パターンの変形処理を説明するための図であ
る。

【図16】図1に示すピッチパターン生成装置の変形例
を示す図である。

【図17】パターン記憶部に記憶されている各種のユニ
ットパターンを示す図である。

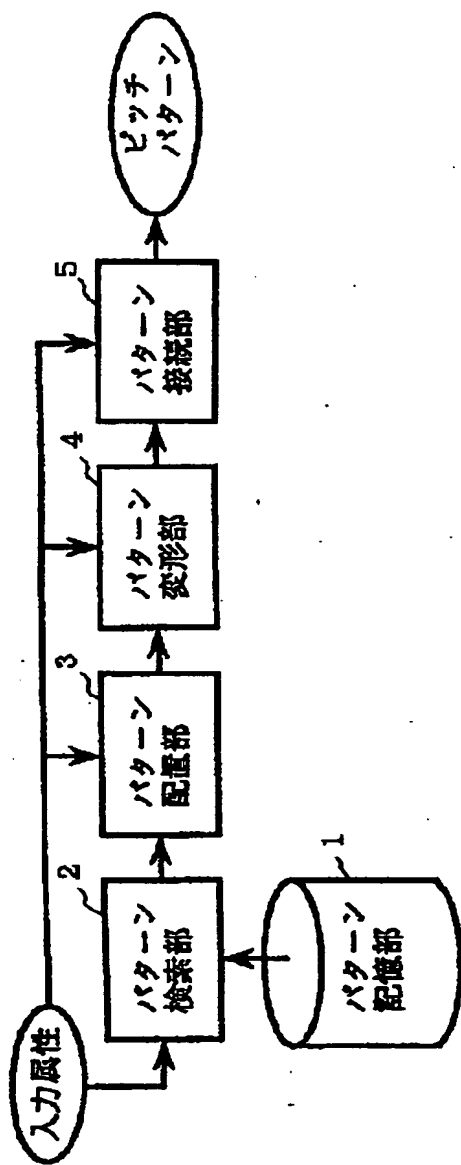
【図18】ルール記憶部に記憶されているルールの構成
例を示す図である。

【図19】入力モーラ列の一例を示す図である。

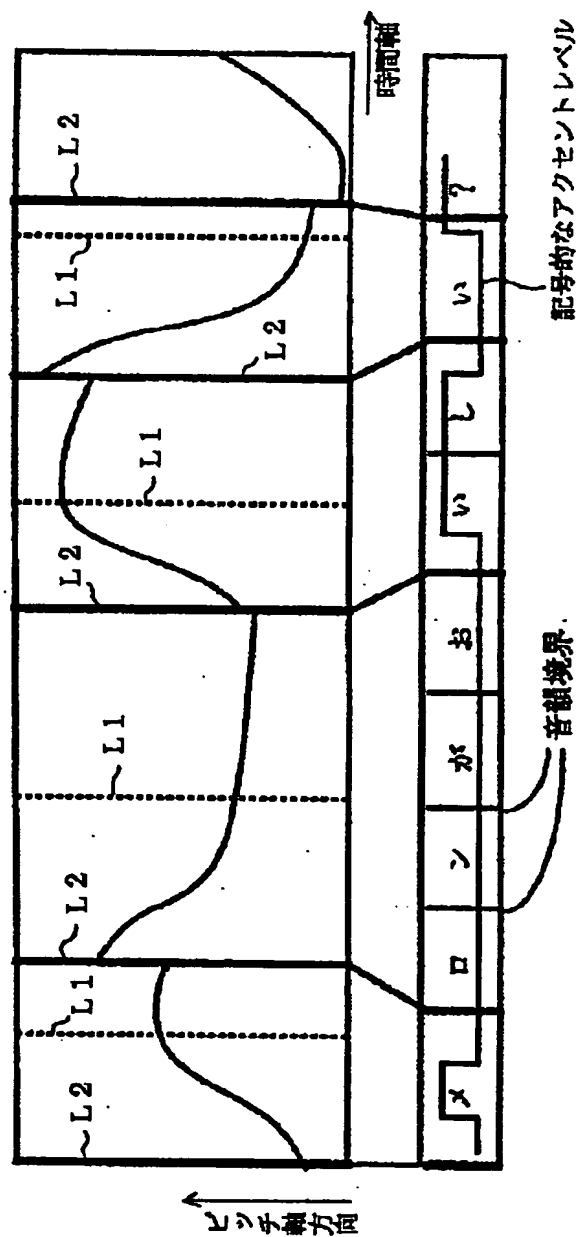
【符号の説明】

- 1 パターン記憶部
- 2 パターン検索部
- 3 パターン配置部
- 4 パターン変形部
- 5 パターン接続部
- 7 ルール記憶部

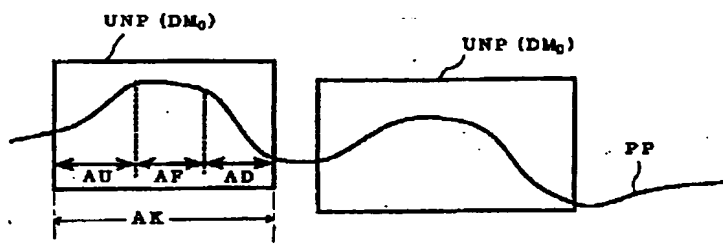
【図1】



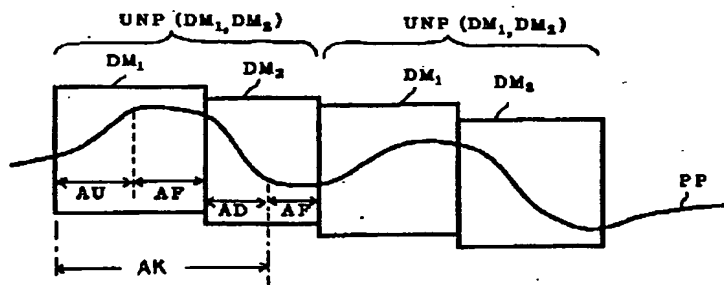
【図7】



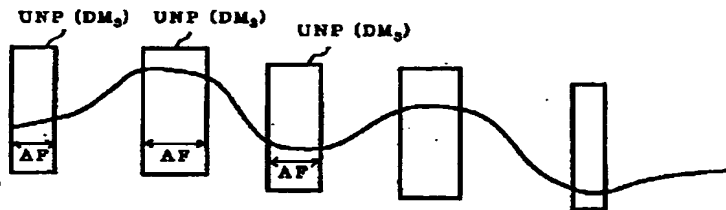
【図2】



【図3】



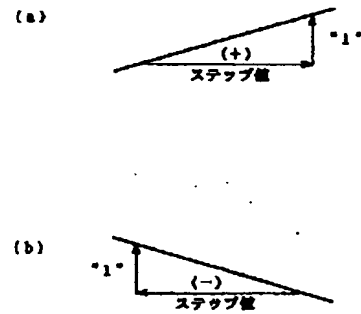
【図4】



【図5】



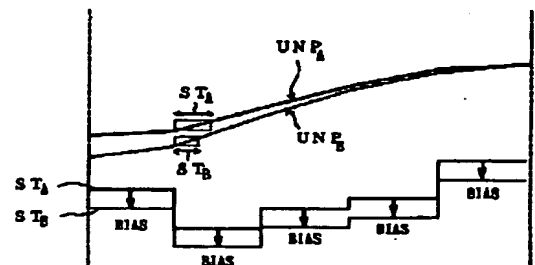
【図11】



【図18】

ユニット番号	"1"	"2"	"3"	"4"
Pos	1	0	1~	1~
Mora	0~5	0~5	0~5	0~5
Acc	0	1	2~	1
Core	0	0	0	~1

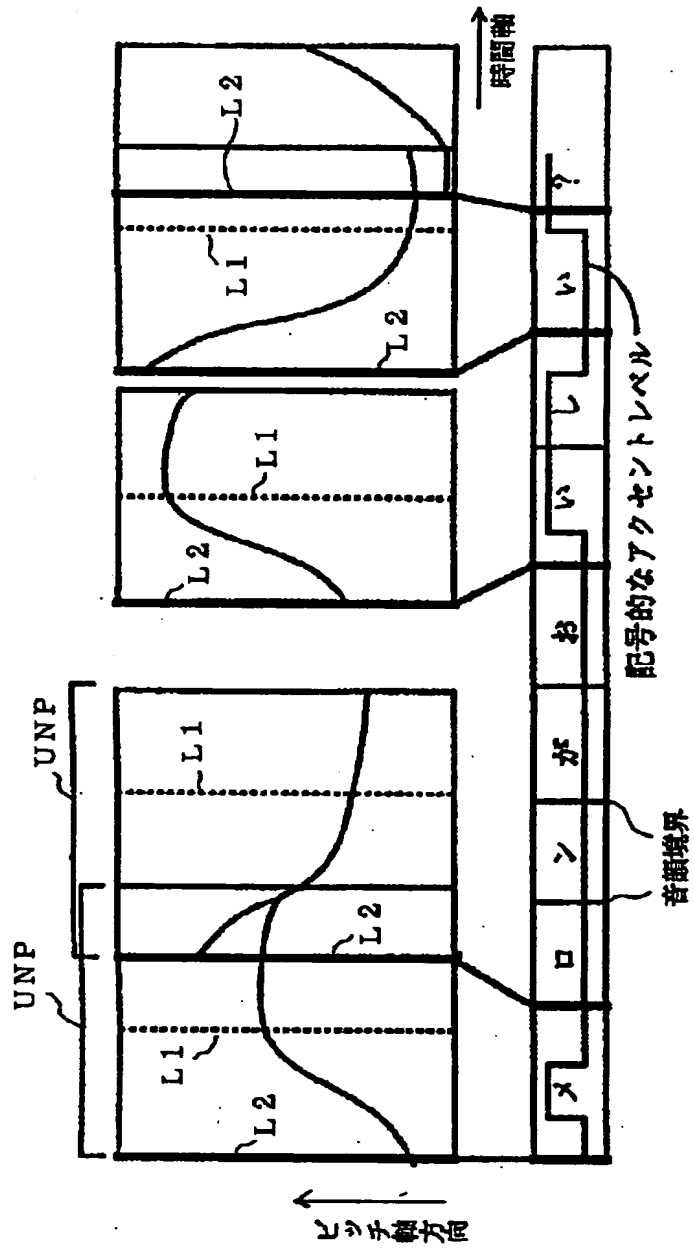
【図13】



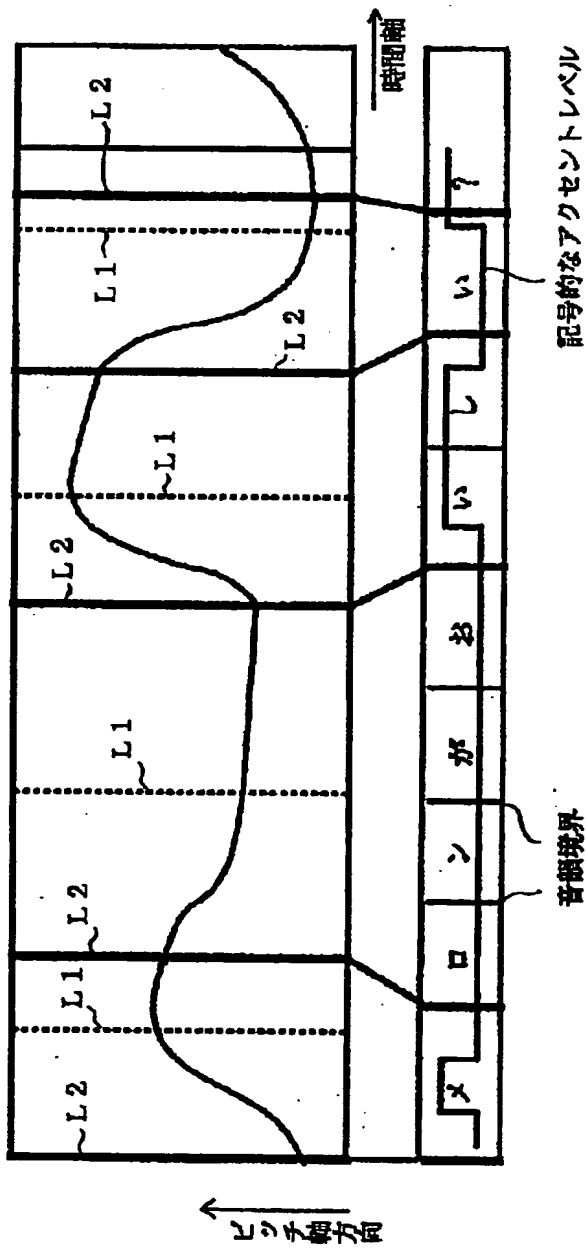
【図14】

ピッチ変化量	バイアス
-0.7~-1.0	+4
-0.5~-0.7	+3
-0.3~-0.5	+2
-0.1~-0.3	+1
-0.1~+0.1	0
+0.1~+0.3	-1
+0.3~+0.5	-2
+0.5~+0.7	-3
+0.7~+1.0	-4

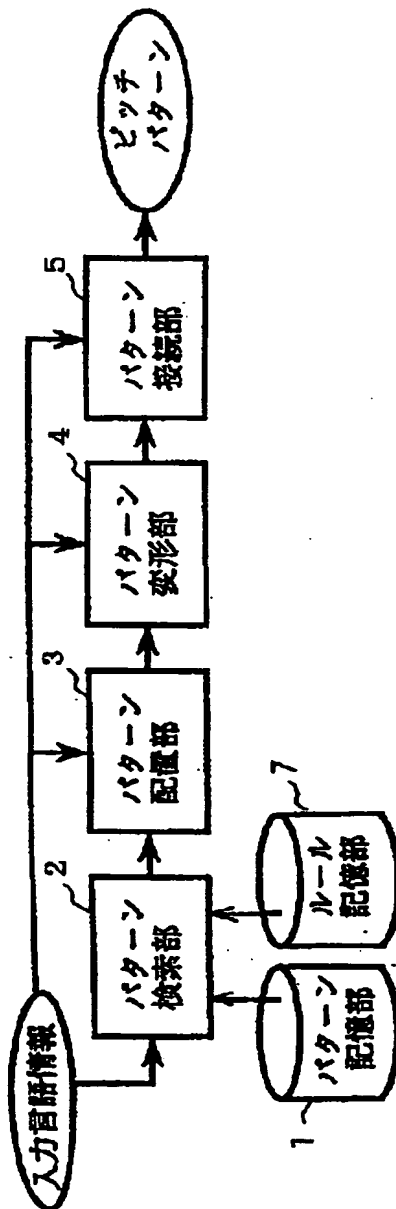
【图 6】



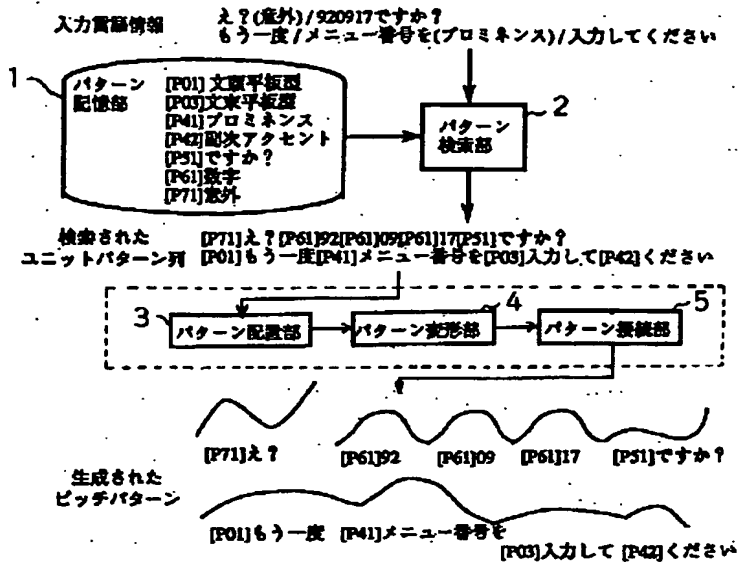
【図 8】



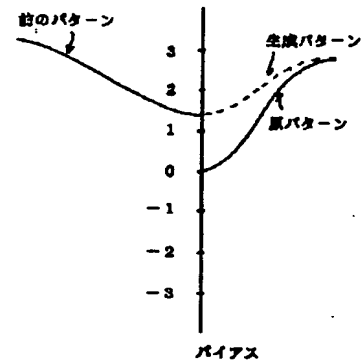
【図 16】



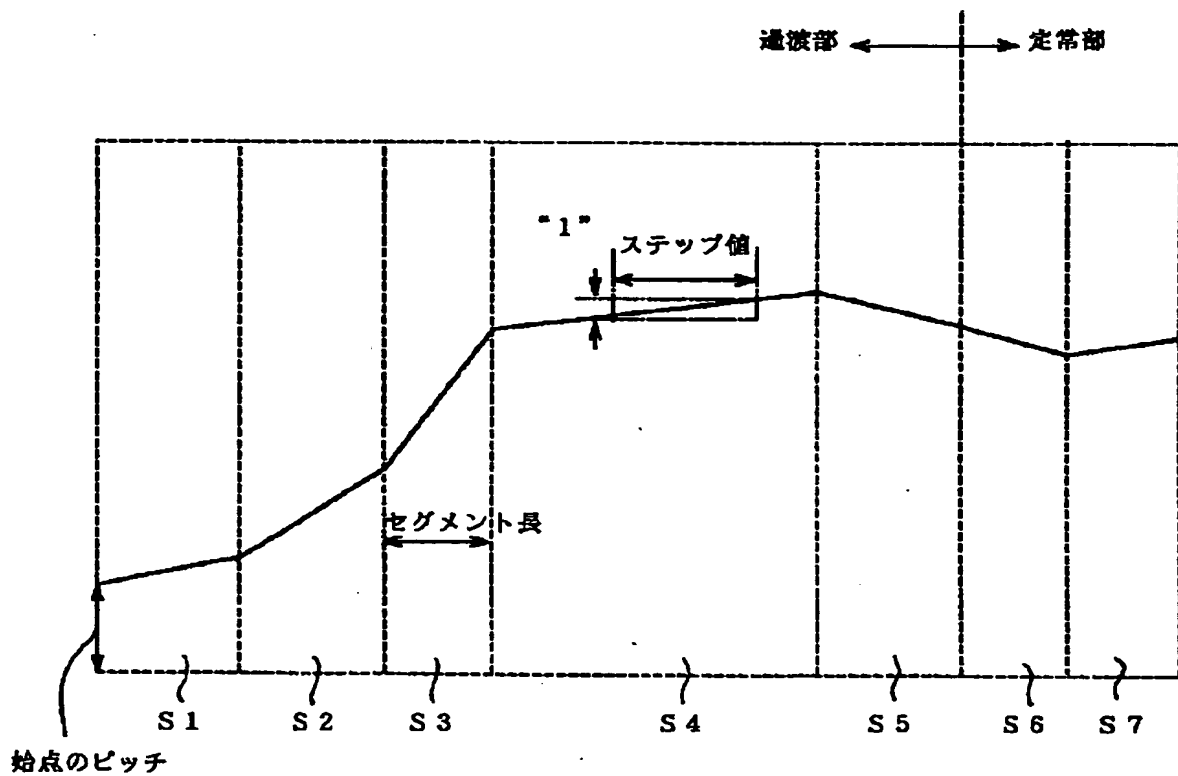
【図9】



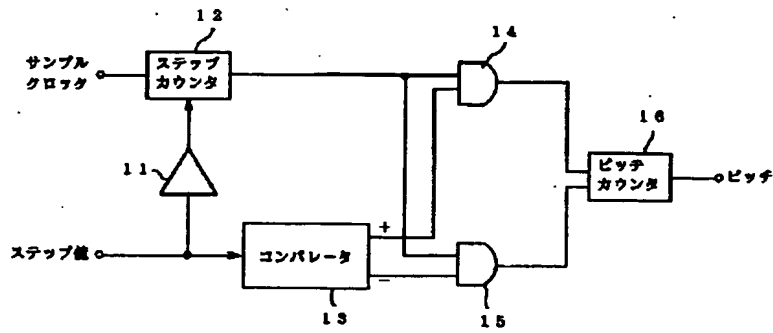
【図15】



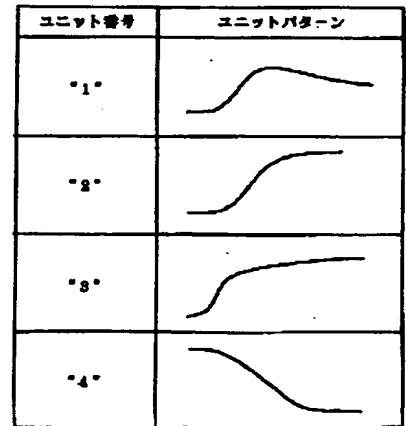
【図10】



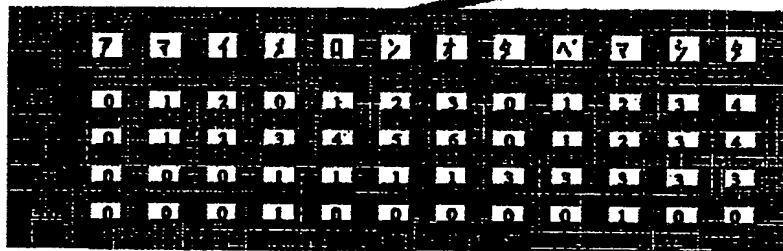
【図12】



【図17】



【図19】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.